PCT/GB 96/02960

The Patent Office

The Patent Office
Cardiff Road
Newport
Gwent
NP9 1RH

RECD 13 FEB 1997
WIPO PCT

PRIORITY DOCUMENT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation and Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 16 DEC 1996

Patents Form 1/77

...its Act 1977
(Rule 16)

The Patent Office

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — P.70330B AHB

2. Patent application number *(The Patent Office will fill in this part)* — -8 JAN 1996 **9600304.1**

3. Full name, address and postcode of the or of each applicant *(underline all surnames)* — The Anthony Nolan Bone Marrow Trust,
The Royal Free Hospital,
Pond Street, Hampstead,
London NW3 2QG.

   Patents ADP number *(if you know it)* — 06943070001

   If the applicant is a corporate body, give the country/state of its incorporation — UNITED KINGDOM

4. Title of the invention — METHODS FOR SEPARATING AND IDENTIFYING DNA MOLECULES

5. Name of your agent *(if you have one)* — J A KEMP & CO

   "Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)* — 14 SOUTH SQUARE
GRAY'S INN
LONDON
WC1R 5LX

   Patents ADP number *(if you know it)* — 06943070001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*
   a) any applicant named in part 3 is not an inventor, or
   b) there is an inventor who is not named as an applicant, or
   c) any named applicant is a corporate body.
   See note (d)) — YES (but will be filed later)

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	33
Claim(s)	7
Abstract	1
Drawing(s)	3

NO

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature [signature] Date 8 January 1996

12. Name and daytime telephone number of person to contact in the United Kingdom

BRASNETT, Adrian Hugh
0171 405 3292

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

METHODS FOR SEPARATING AND IDENTIFYING DNA MOLECULES

The invention relates to methods for separating and identifying DNA molecules in mixtures of DNA molecules having the same number of nucleotides but different base sequences.

1. Background to the Invention

1.1 General Introduction

Genes exist as multiple alleles which differ from each other by small differences in sequence. Individuals are often heterozygous with respect to the alleles of particular genes; i.e. individuals often have two different alleles of the same gene.

In some circumstances, it is desirable to separate the alleles of a gene from a mixture of the alleles. For example, when it is desired to carry out a test to determine which alleles of a gene are carried by a heterozygous individual, it is often necessary to separate the two alleles before carrying out the test because the presence of two alleles in one test can prevent meaningful results from being obtained.

In view of the fact that the difference between the alleles of a gene can be as little as one nucleotide, it is often difficult to separate the alleles from a mixture of the alleles. These difficulties are increased in genes which have a very large number of different alleles, such as the major histocompatibility complex (MHC) genes (e.g. the human leucocyte antigen (HLA) class I genes which have 222 known alleles).

HLA matching between a bone marrow or kidney recipient and donor is one of the major factors influencing transplant

success. Up to date the most favourable bone marrow transplant (BMT) and kidney transplant results have been obtained using sibling donors who are genotypically HLA-identical to the recipient but such donors are available for only about 30% of patients[1-5]. BMT using unrelated donors can be successful, but these transplants have higher rates of graft failure, increased incidence and severity of Graft versus Host Disease and more frequent complications related to delayed or inadequate immune reconstitution [4].

New molecular biological methods for detection of genetic polymorphism currently provide an opportunity to improve e.g. HLA matching of unrelated donors as well as a research tool to investigate the relationship between disparity and transplant complications. These molecular typing methods include sequence-specific amplification, hybridisation with oligonucleotide probes, heteroduplex analysis, single strand conformation polymorphism and direct nucleotide sequencing.

Each of these molecular approaches has been used for routine HLA class II typing [6], but a variety of reasons related to the HLA class I gene structure has complicated their application to class I typing. The reasons for these limitations are the extensive polymorphism of each class I locus and the degree of sequence homology between the loci. In addition, sequence homology between class I classical and non-classical genes and the reported 12 pseudo genes can cause problems for specific locus amplification[7].

The low extent of "allele specific" sequences at polymorphic sites is a feature of the HLA class I genes that has limited the resolution of all current DNA typing approaches. An "allele specific" sequence is a sequence that is only present in one allele and can therefore be used to distinguish the allele from other alleles.

The main problem which complicates the identification of an allele is the presence of a mixture of alleles, as well as contamination by segments of DNA which have homology to the allele it is wished to identify and which are co-amplified in PCR. Current typing methods are sometimes unable to resolve the allele it is wished to identify from contaminating DNA fragments. Separation techniques such as single strand conformation polymorphism (SSCP) can only partially resolve this problem.

1.2 Sequence specific primer amplification (PCR-SSP)

This method utilises both the group-specific and, when present, allele-specific sequence sites in PCR primer design. Although each SSP may not individually define the required specificity, the use of combinations of sequence-specific primers allows the amplification of their common sequences to give the desired HLA specificity. The SSP design is based on the amplification refractory mutation system (ARMS), in which a mismatch at the 3' residue of the primer inhibits non-specific amplification[8,9].

However, despite its high accuracy, PCR-SSP is only in some cases more informative than serology. The reason for this is the low extent of allele specific polymorphic sites in the exon sequences and this limitation has produced a vast amount of research into the identification of allele specific motifs localised in the intron sequences[10]. However, up to date this information has not contributed considerably to the identification of more alleles.

Another limitation of this method is that it detects a limited number of polymorphic sequences which are utilised to predict the entire sequence. If an unknown allele is present in a particular sample this extrapolation may be incorrect.

1.3 Single strand conformation polymorphism (SSCP)

This technique is based on the electrophoretic mobility of single stranded nucleic acids in a non-denaturing polyacrylamide gel, which depends mainly on sequence-related conformation[11-13]. The technique can be employed for isolating single alleles which could then be used for further manipulation and analysis such as direct sequencing. The pattern of bands obtained after electrophoresis may be diagnostic for an allele. Although the technique is easy to perform, the specificity and sensitivity are poor and its use has been limited to only few HLA class I subtypes[14,15].

The successful use of the technique relies on group specific amplification and therefore prior knowledge of broad HLA specificity is needed.

1.4 Heteroduplex analysis

Fully matched DNA duplexes are more stable than those with base mismatches. Instability of the duplex increases with the number of nucleotide mismatches which produce an increasing "drag effect" in polyacrylamide gels which retard the migrating bands[16-18].

This technique exploits the formation of mismatched DNA hybrids (heteroduplex) which are formed at the end of each PCR cycle between coamplified alleles from a particular locus or loci. During the annealing stage of each cycle of the PCR, a proportion of sense strands of each allele may anneal to anti-sense strands of different alleles. A band pattern is obtained in PAGE analysis which can be useful for identifying the alleles involved in the reaction[6,19-21].

Heteroduplex analysis is an approach that has been utilised to compare HLA genes of a particular donor and recipient. HLA genes are amplified, denatured (melted into single

strands) and mixed together under conditions that promote renaturation to form double stranded molecules. If the HLA genes of a donor and recipient are similar but not identical, heteroduplexes will form consisting of one strand of an allele of donor origin and a second strand from a different allele of recipient origin[22,23]. Heteroduplex molecules can be detected by altered electrophoretic migration. The sensitivity of this method can be increased by adding DNA from an HLA allele that is not present in the donor or recipient.

The major advantage of heteroduplex analysis is that it is relatively easy and inexpensive. Limitations of this approach include inability to detect certain HLA disparities, potential detection of irrelevant silent mutations and lack of specific information regarding the nature of the alleles involved.

Up to date this approach has been used only for HLA class II typing with limited success. It has not been applied to HLA class I analysis.

1.5 Sequence specific oligonucleotide probes (PCR-SSO)

SSO typing involves amplification of HLA alleles from a particular locus followed by hybridisation with a panel of oligonucleotide probes to detect polymorphic sequences that distinguish one allele or group of alleles from all others. This is sometimes referred to as low resolution or generic oligotyping. Following low resolution typing, selected primers can be used to achieve amplification of individual alleles which are then identified by specific probes. This second stage of oligotyping is often referred to as high resolution oligotyping[6].

The major methodological drawback of this approach is that the complexity of the technique is directly related to the

number of alleles under investigation and the presence of two alleles in the heterozygous condition can complicate the identification process.

Based on PCR-SSO several typing approaches for HLA-A and B have been published and the number of probes that they require are over 40 and 90 probes respectively[24,25]. This leads to a highly complex operation which is time consuming and the resolution obtained is only moderate.

Published oligotyping methods could result in incorrect interpretation of data if certain combinations of recently discovered alleles are present in a specimen[26]. This limitation and the great number of probes needed have led several groups to explore the feasibility of automated sequencing for routine HLA class I typing.

The advantages of PCR-SSO methods are specificity, sensitivity, simplicity, reproducibility, inexpensive and that they allow simultaneous processing of many samples.

1.6 Direct Nucleotide Sequencing

This technique involves PCR amplification using one biotinylated and one non-biotinylated primer. Following amplification, PCR products are captured by streptavidin-coated magnetic beads and the non-biotinylated strand is removed, this permits unidirectional solid-phase sequencing of purified single-stranded DNA[6].

Sequencing methods can be differentiated by whether the template has been cloned or produced directly from genomic DNA by an initial PCR amplification. Sequencing approaches can also employ either manual or automated techniques[14]. Cloned templates and templates produced from allele-specific PCR amplification represent a single sequence derived from one HLA haplotype. In contrast, alleles from

both HLA haplotype in heterozygous samples may be co-amplified and sequenced together using locus-specific PCR primer. The recent availability of software, which allows the user to align the derived sequence against established libraries, has facilitated the analysis and allele assignments for heterozygous samples in which both templates are sequenced at the same time[27]. The effectiveness of this method depends on the amount and frequency of ambiguous heterozygous combinations which for HLA class II are not many, but for class I the number of ambiguous combinations is so high that the obtained resolution is comparable with serology.

Up to date two HLA class I typing approaches based on direct sequencing have been published, both require serology information followed by allele specific PCR amplification and then direct sequencing[14,27]. Although this approach gives the highest resolution, it is expensive and difficult to perform routinely.

2. Summary of the Invention

The invention provides methods for separating and identifying a DNA molecule in a mixture of DNA molecules having the same number of neucleotides but different base sequences. The separation methods comprise

(i) amplifying the DNA molecules in the mixture;

(ii) hybridising single strands of the amplified DNA molecules with a complementary strand of a reference DNA molecule so as to form duplexes; and

(iii) separating the duplexes.

The different DNA molecules in the original mixture give

rise to duplexes having different numbers, positions or types of mismatches. This allows the duplexes to be separated by, for example, gel electrophoresis. The separated duplexes can then be analysed to identify the DNA molecules that were present in the original mixture. One embodiment of a method of the invention is illustrated in Figure 1.

The method of the invention can be used directly as a diagnostic technique to identify a DNA molecule by the use of a specific reference DNA molecule. The formation of a homoduplex identifies a DNA molecule in the unknown mixture as identical to the reference DNA molecule. The formation of a heteroduplex may also be used to identify an unknown DNA molecule by using a known heteroduplex as a control.

The method can also be used as a separation technique for separating the alleles in a mixture of unknown alleles of a polyallelic gene, such as the mammalian MHC genes (e.g. the HLA genes). Duplexes formed between the unknown alleles and a reference allele are separated so as to isolate the unknown alleles for identification by techniques such as DNA sequencing, SSP and SSO.

In one embodiment, the invention provides a method which comprises

(i) amplifying the DNA molecules in the mixture employing a pair of primers in which one of the primers carries a ligand, so as to produce an amplified mixture of double-stranded DNA molecules in which one of the strands carries a ligand;

(ii) contacting the amplified mixture of double-stranded DNA molecules with a receptor on a solid support under conditions such that the ligand

binds to the receptor;

(iii) separating the mixture of double-stranded DNA molecules into single-strands and removing the strands that are not bound to the support by the ligand;

(iv) recovering the remaining strands from the support;

(v) mixing the recovered strands with a complementary strand of a reference DNA molecule so as to form duplexes; and

(vi) separating the duplexes.

The complementary strand of the reference DNA molecule may be provided by essentially the same technique as the technique set out above in steps (i) to (iv) for providing the mixture of DNA molecules in single-standard form. In particular, the reference complementary strand DNA molecule may be provided by

(i) amplifying the reference DNA molecule employing a pair of primers in which one of the primers carries a ligand, so as to produce amplified double-stranded reference DNA molecule in which one of the strands carries a ligand;

(ii) contacting the double-stranded reference DNA molecule with a receptor on a solid support under conditions such that the ligand binds to the receptor;

(iii) separating the double-stranded reference DNA molecule into single-strands and removing the

strand that is not bound to the support by the ligand; and

(iv) recovering the remaining strand from the support.

In another embodiment of the invention, there is provided a method which comprises

(i) amplifying the DNA molecules in the mixture employing a pair of primers in which one of the primers carries a high molecular weight molecule, so as to produce an amplified mixture of double-stranded DNA molecules in which one of the strands carries a high molecular weight molecule;

(ii) separating the mixture of double-stranded DNA molecules into single strands;

(iii) mixing the single strands with a complementary strand of a reference DNA molecule so as to form duplexes; and

(iv) separating the duplexes.

The complementary strand of the reference DNA molecule may be provided by

(i) amplifying the reference DNA molecule employing a pair of primers in which one of the primers carries a high molecular weight molecule, so as to produce an amplified double-stranded reference DNA molecule in which one of the strands carries a high molecular weight molecule; and

(ii) separating the double-stranded reference DNA

molecule into single strands.

This embodiment overcomes the need for solid support systems by conjugating one primer of a pair of primers directly to a high molecular weight molecule (e.g. a protein), for the reference and test systems. The amplified product after hybridisation can be applied directly to a separating gel. The high molecular weight conjugates are retained in the gel compared to the duplex without attachment of the high molecular weight molecule.

In a further embodiment of the invention, there is provided a method which comprises

(i) amplifying a single strand of each of the DNA molecules in the mixture;

(ii) mixing the amplified single strands with a complementary strand of a reference DNA molecule so as to form duplexes; and

(iii) separating the duplexes.

In this embodiment, the complementary strand of the reference DNA molecule may be provided by amplifying a single strand of the reference DNA molecule. The amplification of the single strand of the reference or test DNA molecule can be done, for example, by asymmetric PCR.

This embodiment overcomes the need for both solid support systems and conjugation of one primer of a pair to a high molecular weight molecule. However, in the embodiment it is possible to use a primer carrying a ligand such as a hapten in order to facilitate capture of the amplified strand with a receptor such as an antibody and separation of the amplified strand from other components in the

amplification mixture.

After separation of the DNA molecules by one of the above methods, the molecules present in the mixture may be identified by carrying out one or more of the following steps:

(i) comparing the positions of the separated duplexes on the gel with the position of a control DNA;

(ii) sequencing each of the separated molecules;

(iii) sequence specific primer (SSP) amplification analysis; and

(iv) sequence specific oligonucleotide (SSO) analysis.

The invention provides an improvement over prior methods for separating DNA molecules. The advantages offered by the invention can be summarised as follows:

(a) The invention provides a high resolution between different DNA molecules and differences of as little as one nucleotide between molecules can be detected.

(b) The invention allows simultaneous and rapid processing of a large number of samples.

(c) The invention is comparatively inexpensive to perform, particularly when compared to prior methods which achieve a high level of resolution.

(d) The invention uses techniques that can be performed easily without recourse to complex and expensive technology.

3. Principle underlying the invention

Fully matched DNA duplexes are more stable than those with base mismatches. Regions of nucleotide sequence complementary are double stranded, but mismatched regions form single-stranded loops along the length of the DNA molecule. The number, size, composition and position of the single-stranded loops vary for each combination of alleles. The rate at which the DNA migrates in polyacrylamide or special agarose gels depends on both molecular conformation and molecular weight. Heteroduplex DNA always migrates more slowly than the corresponding homoduplex DNA. Both denaturing reagents and/or heat enhance the degree of separation of the mismatched DNA strands (heteroduplexes) from the corresponding matched duplexes (homoduplexes).

As the molecular conformation of heteroduplexes can be manipulated by hybridisation of a known single strand reference with unknown complementary single strand(s), it is proposed that heteroduplexes can be separated from each other by e.g. denaturing or non-denaturing polyacrylamide electrophoretic analysis. This allows the separation of the two amplified alleles from a particular locus for further analysis. In addition, the method of the invention (which we call "Complementary Strands Analysis" (CSA)), permits assessment of the quality of the PCR product before the process of identification is carried out. CSA is able to identify the presence of coamplified non-desirable alleles from different loci and, potentially, PCR fragments that contain artifacts such as Taq errors and in vitro recombinations.

In addition, CSA itself can be used as a diagnostic technique. It can identify alleles by hybridisation of allele specific single strand with unknown complementary single strand(s) followed by e.g polyacrylamide gel

electrophoretic analysis, with or without denaturing conditions and/or with or without a temperature gradient. The formation of a homoduplex demonstrates identity between at least one of the unknown alleles and the allele specific reference, and non-identity between them produces heteroduplex(es).

4. Detailed Description of the Invention

The kinds of DNA molecule that may be separated and identified by the methods of the invention include alleles of polyallelic genes, segments of genes and non-expressed fragments.

Examples of genes with multiple alleles to which the invention may be applied are the mammalian MHC genes such as the HLA class I and class II genes, the T cell receptor genes in mammals [30,31], TAP, LMP, ras[29], non classical HLA class I genes, the genes for human complement factors C4 and C2, Bf in the human HLA complex, and genes located in mitochondrial DNA, bacterial chromosomes and viral DNA. The invention can be used in the analysis and identification of mutations (e.g. point mutations) in these and other genes and chromosomal aberrations such as translocations, deletions and inversions.

There are three different genes within the HLA class I group of genes, namely HLA-A, HLA-B and HLA-C, and each of these three genes exists in the form of multiple alleles. There are a total of about 222 known alleles of the HLA-A, HLA-B and HLA-C genes and the sequences of known alleles are set out in Arnett and Parham (1995) Tissue Antigens 45 217-257. There are also multiple genes within the HLA class II group of genes, known as DR, DQ and DP.

In the method of the invention, it is necessary to identify primer sequences unique for the target gene so as to

include all polymorphic sites of interest in the amplified fragment, which should also be manageable in length. For example, the polymorphic sites in exons 2 and 3 of HLA class I would facilitate the identification of all recognised alleles of HLA-A, B and C, with 5 exceptions, if locus-specific amplification is achieved. Therefore, the primers used in the invention may, for example, be selected so as to specifically amplify exons 2 and 3 of each of HLA-A, HLA-B and HLA-C separately. Cereb[28] and collaborators have described primer sequences located in the first and third exons which can be used for locus-specific amplification of the entire exon 2 and 3 region of each of the HLA-A, HLA-B and HLA-C genes. The sequences of suitable primers are given in the Example below.

The reference DNA molecule used in the invention generally has a known sequence. The reference may be chosen so as to have a similar allotype to an allotype that at least one of the test alleles is suspected of having. For example, it may be known that a test allele is of the HLA-A02 type from serological data, but it may not be known which of the seventeen A02 sub-types the allele is. In this case, the reference allele may be chosen to be of sub-type A0201 and the method of the present invention could then be used to determine which of the A02 sub-types the test allele is.

The reference strand may be obtained from (a) a homozygous source, (b) a heterozygous source from which individual strands are isolated by gel separation after amplification steps or (c) DNA synthesis. There are now about 500 internationally recognised cell lines which contain HLA alleles of known sub-type and these cell lines can be used as a source of reference alleles.

The initial broad typing of the test allele may be performed using serological techniques. Different alleles of HLA induce different types of antibody and the antibody

induced by an unknown HLA therefore gives information as to the allotype of the HLA. However, serological typing does not achieve a high level of resolution and is not sufficiently accurate or reproducible for matching a prospective donor with a prospective recipient in a tissue or organ transplant operation.

The control DNA used in the method of the invention may be a homoduplex between two strands of the same DNA molecule (e.g. the reference DNA molecule), so that migration of a test duplex to the same position on the electrophoretic gel as the control homoduplex indicates that the test duplex is a homoduplex. If the test duplex is a homoduplex, it can be concluded that the unknown DNA molecule is the same as the reference molecule.

Control DNAs may be obtained by simply amplifying a known DNA molecule using the same primers as used in the method of the invention to amplify the reference and unknown molecules.

The control DNA may also be a heteroduplex of known DNA molecules. This allows the method of the invention to be used to identify molecules in heteroduplexes formed by test samples. The same heteroduplexes from different sources migrate to the same position on a gel.

A potential problem with identifying molecules in heteroduplexes is that certain different heteroduplexes containing the same number and type of mismatches migrate to the same position. However, this problem can be overcome. For example, in typing of the HLA class I alleles, the different duplexes could be identified by their different sizes (the HLA-A, HLA-B and HLA-C genes are different sizes) or by amplifying each of the HLA-A, HLA-B and HLA-C genes with primers carrying different labels. Each locus specific heteroduplex would have a different

size or carry a different label, and could be electrophoresed simultaneously in the same track of a gel. The duplexes could then be identified by comparing them to control duplexes in the same gel. Examples of suitable labels include radiolabels, colour labels and fluorescent labels.

The mixture of alleles used in the method of the invention may be from a prospective donor or a prospective recipient in a tissue or organ transplant operation. The results of the method may therefore be used to match a prospective recipient with a prospective donor.

In one embodiment of the invention, the alleles of the prospective donor or of the prospective recipient are in effect used as reference alleles and duplexes are formed between strands of the prospective recipient's alleles and of the prospective donor's alleles. Analysis of the duplexes formed between the strands from the prospective recipient and donor reveals whether they have compatible alleles. Thus, in one embodiment, the invention provides a method for determining whether a prospective recipient in a tissue or organ transplant operation has alleles of a gene that are compatible with the alleles of a prospective donor in the operation, which method comprises

(i) amplifying the alleles of the prospective recipient employing a pair of primers in which one of the primers carries a ligand, so as to produce amplified double-stranded alleles of the prospective recipient in which one of the strands carries a ligand;

(ii) contacting the amplified double-stranded alleles with a receptor on a solid support under conditions such that the ligand binds to the receptor;

(iii) separating the double-stranded alleles into single-strands and removing the strands that are not bound to the support by the ligand;

(iv) recovering the remaining strands from the support;

(v) mixing the recovered strands with complementary strands of the alleles of the prospective donor so as to form test duplexes;

(vi) separating the test duplexes by gel electrophoresis; and carrying out one or more of the following steps:

(vii) comparing the positions to which the test duplexes migrate on the gel with the position of a control DNA;

(viii) sequencing one or both strands of each of the test duplexes;

(ix) sequence specific primer (SSP) amplification analysis; and

(x) sequence specific oligonucleotide (SSO) analysis.

Other proposed uses of the invention include determination of the paternity of an individual by identifying one (or more) of his alleles to see if it is the same as a corresponding allele of a potential father. The invention may also be used in forensic medicine to determine the origin of a sample of body tissue or fluid, as a follow up technique in treatment of haematological malignancies or inherited disorders, in adoptive immunotherapy, and in identification of bacteria and viruses.

In the method of the invention, the amplification steps may be carried out by polymerase chain reaction (PCR).

The ligand/receptor system used in the invention may, for example, be the biotin/streptavidin system or a hapten/antibody system. Direct conjugation of the primer via a linking group, such as short poly A, to the beads is an alternative. When the biotin/streptavidin system is used, one of the primers used in each of the amplification steps may be labelled with biotin, so that when the amplification reaction is carried out double-stranded DNA is produced in which one strand carries a biotin label. The double-stranded DNA may then be bound to a solid support coated with streptavidin.

The solid support used in the invention is typically magnetic beads. However, other supports may be used, such as the matrix of an affinity chromatography column. When the support is in the form of magnetic beads, the two strands of the amplified DNA are separated by attracting the beads to a magnet and washing the beads under conditions such that the double-stranded DNA dissociates into single-strands. The dissociation is typically performed by incubating the beads three times under alkaline conditions (e.g. 0.1 M NaOH) at room temperature for about 5 minutes. Usually, the strand which is not bound to the support by the ligand is then discarded, although it is equally possible to retain the strand that is not bound to the support and discard the strand that is bound to the support.

The strand that remains attached to the support may be recovered from the support by incubating the support under conditions such that the ligand/receptor complex dissociates. When the biotin/streptavidin system is used, the support is typically heated to e.g. 95°C for about 5 minutes; this ensures denaturation of the streptavidin

molecule to release the biotinylated single strand which is then recovered.

At this stage, there have been provided a single-stranded unknown allele and the complementary strand of a reference allele. The two strands are then mixed together under conditions in which they hybridise to form duplexes. Typically, the hybridisation step is performed by heating the mixture of strands at about 95°C for about 3 min, at about 70°C for about 5 min and then at about 65°C for about 45 min.

Under these conditions, duplexes are formed which can subsequently be separated by gel electrophoresis (e.g. polyacrylamide gel electrophoresis). The electrophoresis is preferably carried out under denaturing conditions because this has the effect of amplifying the loops formed by mismatched strands and hence allows better resolution.

As an alternative separation technique to gel electrophoresis, high pressure liquid chromatography (HPLC) may be used.

In the embodiment of the invention in which one of the pair or primers is conjugated to a high molecular weight molecule, the molecule may be a protein such as bovine serum albumin (BSA). The molecular weight of the high molecular weight molecule is such that it causes the DNA molecule to which it is attached to be sufficiently retarded in the separation step (e.g. the electrophoresis step) to allow the DNA molecule to be separated from a duplex without a high molecular weight compound attached. For example, the molecular weight of the high molecular weight molecule may be from 10 to 200 kDa, preferably 20 to 100 kDa.

<u>Use of the invention to analyse ras oncogene point</u>

mutations:- ras has been implicated in the oncogenesis of many tumours and appears to be activated by point mutations. These mutations can occur in all three ras genes (N-ras, Harvey-ras and Kirsten-ras) at codons 12/13 and 61 with corresponding amino acid substitutions in ras proteins (p21). These point mutations can be detected by application of the invention.

Two pairs of primers are needed, one for the 12/13 codons and one for codon 61. The primers described by Lyons [29] can be used, with modification by covalent attachment of a ligand to one primer of each pair for each of the test fragment and the reference fragment. For example, a ligand may be attached to primer P1a (a 12/13 codon primer) and primer P1b (a 61 codon primer) for the reference fragment, and to primer P2a (a 12/13 codon primer) and primer P2b (a 61 codon primer) for the test fragment. In this way, complementary ligand-labelled single strands for the reference and test fragments are obtained. The complementary strands are hybridised and subjected to electrophoresis. Detection of a homoduplex between the test fragment and a mutant reference fragment will indicate the presence of the mutation in the test fragment.

Use of the invention to identify T cell receptor (TCR) rearrangements in T cell tumours and in adoptive immunotherapy:- some T cell tumours can be monoclonal in origin and a proportion of the T cells from a patient may carry a particular rearrangement of the T cell variable domain genes alpha/beta or gamma/delta depending on the T cell type. The efficacy of a particular treatment or the course of the disease can be evaluated by the identification of the malignant clone TCR rearrangement. The method of the invention with the use of suitable number of controls can be made semi-quantative, which would allow the evaluation of the progress of the treatment or the disease.

In adoptive immunotherapy, a specific rearrangement of the variable domain genes of the TCR can be used as a marker for the selected cytotoxic T cell that has been generated in-vitro. Post infusion fate of these cells can be monitored by a semi-quantative detection of the particular rearrangement.

In both methods the T cell variable domain gene primers [30,31] can be modified by covalent attachment of a ligand at the 5′ end of the primer pair, and the reference strand will be selected to be complementary to the test DNA fragments.

EXAMPLE

The following Example illustrates the invention.

1. Brief Description of the Drawings

Figure 1 shows a schematic overview of an embodiment of the method of the invention.

Figure 2 shows the band pattern formed between the sense strand of LO541265 (A*0101) with the anti-sense strand(s) of STEINLIN (A*0101), WIN (A1), DAUDI (A*0102, A*6601), BM21 (A1) and E4181324 (A1) under the following conditions: 8% polyacrylamide, 2 M urea, 10% formamide, 200 Volts during 6 hrs at room temperature.

Figure 3 shows duplexes formed between the sense strand of JY (A*0201) with the anti-sense strand(s) of M7 (A*0202, A*0301), DK1 (A*0203, A*3301), RML (A*0204), WT49 (A*0205), CLA (A*0206, A24), KNE (A*0207, A*0201), KLO (A*0208, A*0101) and OZB (A*0209, A*0301) under the following conditions: 8% polyacrylamide, 2 M urea, 10% formamide, 200 Volts, 6 hrs, at room temperature.

Figure 4 shows the position of a heteroduplex band compared to a homoduplex control. The heteroduplex band corresponds to the hybridisation between the sense strand of JY (A*0201) and the anti-sense strand of RML (A*0204) which differ from each other ny only one nucleotide. Denaturing conditions were: 8% polyacrylamide, 3 M urea, 15% formamide, 200 Volts, 6 hrs. at 50°C.

Figure 5 shows a Complementary Strands Analysis PAG/agarose gel. Following HLA-A locus specific amplification of exons 2 & 3, the antisense DNA strands were isolated and hybridised with complementary locus reference strand. Samples were applied to PAG/agarose cassette and electrophoresis was performed at 230 volts for 6 hrs at room temperature; the gel was stained with SYBR Green I. Single bands were obtained from homozygous cell lines and double bands from heterozygous lines. Duplexes of identical alleles from different sources have the same mobility in the gel; the samples in lanes 5 and 8 are from alleles A0101 and A0201, and the samples in lanes 1 and 9 are from a homozygous line with A0101 which has the same mobility as the fast lines in lanes 5 and 8. Lane 1: STEINLIN (A*0101), Lane 2: KIME (A*0211 - A*3201), Lane 3: DAUDI (A*0102 - A*6601), Lane 4: EA (A*0301), Lane 5: LCL721 (A*0101 - A*0201), Lane 6: M7 (A*0202 - A0301), Lane 7: CJO-A (A*1101), Lane 8: T5-1 (A0101 - A*0201), Lane 9: LO541265 (A*0101), Lane 10: AM (A*0205 - A3201).

2. Methods

2.1 Summary

Amplification and isolation of the biotinylated anti-sense strand(s) were performed. They were then hybridised with allele specific reference sense single strands.

In order to test the resolving power of this method, 7

allele specific reference sense single strands were prepared. These were hybridised with several isolated anti-sense strand HLA-A alleles which were selected to include alleles with one or several nucleotide differences compared to the reference strand.

Following this step, PAGE analysis was performed under several different denaturing conditions. Urea at a range of 2-6 M and formamide between 10%-30% concentrations were used and the gels were run at either room temperature or 50-58°C (200 Volts for 6 hours).

The identification of a homoduplex in the denaturing PAGE gels indicated identity between the allele specific single reference strand and at least one of the alleles from the unknown DNA sample.

As a positive control, double strand DNA PCR product from the allele reference was always used.

2.2 Locus specific amplification of HLA class I genes

For typing purposes amplification of exons 2 and 3 is desirable, and the primers were therefore selected to amplify the stretch of the genome between intron 1 and intron 3. The localisation and nucleotide sequences of the HLA locus-specific primers used are given in the reagents section.

PCR reactions were performed in a total volume of 100µl using 1µg of genomic DNA and 25 pmoles of each locus-specific primer. The 3'-primer was biotinylated at 5'-end. This arrangement ensures the incorporation of the biotinylated primer onto the amplified antisense DNA strand. PCR conditions are given in the following table.

Thermocycling conditions

A and C loci	95°C	4 min.	1 cycle
	95°C	30 sec.	
	70°C	50 sec.	33 cycles
	72°C	30 sec.	
	72°C	8 min.	1 cycle
B locus	95°C	4 min.	1 cycle
	95°C	30 sec.	
	65°C	50 sec	33 cycles
	72°C	30 sec	
	72°C	8 min.	1 cycle

2.3 Separation of the amplified DNA strands

2.3a Removal of non-biotinylated strand:

Magnetic beads with covalently coupled streptavidin on the surface were added to the PCR product and incubated for 30 minutes at 43°C. In this way the amplified PCR product was immobilised by the interaction of biotin and streptavidin. After incubation, the tubes were placed against a magnet and the beads were washed with washing buffer to remove the remaining PCR reaction components.

The non-biotinylated DNA strand was then dissociated from the beads by incubation with 0.1 M NaOH at room temperature (r.t.) for 5 minutes (X3). Following this the beads were washed to remove excess NaOH and resuspended in 50 μl of hybridisation buffer.

2.3b Removal of biotinylated DNA strand:

The bead suspension was heated at 95°C for 5 minutes; this ensures denaturation of the streptavidin molecule to release the biotinylated amplified anti-sense single strand which was then removed and placed in a clean tube. At this stage, the isolates contained single biotinylated DNA strands from each allele.

2.4 Preparation of allele specific reference single-stranded DNA

DNA was extracted from 10th IHW cell lines. The following homozygous cell lines were selected as allele specific reference DNA: JY (A*0201), RML (A*0204), WT49 (A*0205), T7527 (A*0206), KRC-005 (A*0212), AMALA (A*0217) and LO541265 (A*0101).

The PCR conditions for amplification were as above, with the exception that in each case the locus-specific 5'-primer was biotinylated (5'-end). The PCR products were analysed by PAGE to assess the fidelity of the amplification and in all cases a single band was obtained.

2.5 Hybridisation

The biotinylated anti-sense strand(s) from above were mixed with the sense strands, and the mixture was heated at 95°C for 3 min., incubated at 70°C for 5 min., and then at 65°C for 45 min. Under these conditions, the sense and anti-sense strands were hybridised. The heteroduplexes formed could subsequently be separated from each other by electrophoresis in polyacrylamide gel.

3. Reagents:

A) Nucleotide sequences of primers used for locus-specific

amplification:

5' A locus primer: GAA ACG/C GCC TCT GT/CG GGG AGA AGC AA
(Intron 1: 21-46)

3' A locus primer: TGT TGG TCC CAA TTG TCT CCC CTC
(Intron 3: 66-89)

5' B locus primer: GGG AGG AGC GAG GGG ACC G/CCA G
(Intron 1: 36-57)

3' B locus primer: GGA GGC CAT CCC CGG CGA CCT AT
(Intron 3: 37-59)

5' C locus primer: AGC GAG GG/TG CCC GCC CGG CGA
(Intron 1: 42- 61)

3' C locus primer: GGA GAT GGG GAA GGC TCC CCA CT
(Intron 3: 12-35)

B) Buffers:

Washing buffer:	10 mM	Tris-HCl pH 7.5
	1.0 mM	EDTA
	2.0 M	NaCl
Hybridisation buffer:	20 mM	Tris-HCl pH 8.4
	50 mM	KCl
PCR buffer:	20 mM	Tris-HCl pH 8.4
	50 mM	KCl
	0.2 mM	MgCl2
TE buffer	10 mM	Tris-HCl pH 7.5
	1 mM	EDTA

C) Various

Dynabeads M-280 Streptavidin (10 mg/ml)

Magnetic particle concentrator -Dynal MPC

A Thermal cycler (PTC-200 Peltier Thermal Cycler MJ Research)

Ultrapure dNTP set, 2' -Deoxynucleoside 5' -Triphosphate (Pharmacia Biotech)

Taq DNA Polymerase (Gibco BRL)

50 mM MgCl2

0.1 M NaOH

SeaPlaque Agarose (Flowgen Instruments Ltd)

Protogel, 30% Acrylamide and 0.8% Bisacrylamide (National Diagnostics)

4. Results: Identification of alleles by CSA

Seven allele specific reference sense strands were isolated from homozygous cell lines and hybridised to several anti-sense strands from other cell lines whose HLA specificity was defined by sequencing and in some cases only by serology.

Alleles could be identified by homoduplex formation with a reference strand, which would migrate at the same rate as the double stranded DNA reference band (control band) in denaturing polyacrylamide gel. Non-identity would lead to heteroduplex formation and cathodic band(s) compared with the control band.

Under the denaturing conditions of 2 M Urea and 10% formamide

at room temperature, it was observed that when the reference strand matched completely the anti-sense strand a single homoduplex band was visible in the gel. In cases when the reference strand differed by 3 or more bases from the anti-sense strand, a cathodic band corresponding to a heteroduplex was seen. This pattern was observed reproducibly for all the allele specific reference strands and is therefore independent of the position of the mismatch(es) on the strands and the specific base sequence of the allelic reference (Figures 2 and 3).

Under these conditions, the heteroduplex band containing 1 or 2 base mismatches was indistinguishable from the homoduplex band. By altering the denaturing conditions to 3 M urea and 15% formamide at 50°C, it was possible to distinguish between heteroduplexes with one or two mismatches from the homoduplex control band (Figure 4).

In further experiments, it was shown that heteroduplexes of the same alleles from different sources migrate to the same position on the gel. Figure 5 shows the results from such an experiment. Lanes 1, 5 and 8 contain duplexes comprising A0101 from different sources and the duplexes all migrated to the same position. Lanes 5 and 8 also contain duplexes comprising A0201 from different sources and these migrated at the same speed, but more slowly than the A0101 duplexes.

References

1.- D. Weisdorf, R. Haake & B. Blazar. Risk factors for acute graft-versus-host disease in histocompatible donor bone marrow transplantation. *Transplantation* 1991: 51: 1197-1203

2.- L. A. Ochs, W. J. Miller, A. H. Filipovich, R. J. Haake, P. B. McGlave, B. R. Blazar, N. K. C. Ramsay, J. H. Kersey & D. J. Weisdorf. Predictive factors for chronic graft-versus-host disease after histocompatible sibling donor bone marrow

transplantation. *Bone Marrow Transplant* 1994: 13; 455-460

3.- L. A Smyth, C. S. Witt, F. T. Christiansen, R. P. Hermann, P. N. Hollingsworth, D. C. Townend, E Edward & R. L. Dawking. The MHC influences acute graft versus host disease in MHC matched adults undergoing allogeneic one marrow transplantation. *Bone Marrow Transplant* 1993: 12; 351-355

4.- A. Bacigalupo, F. Gualandi, M. T. Van Lint, M. Sessarego, F. Frassoni, D. Occhini, T. Lamparelli, R. Oneto, V. Vitale, R. Corvo, E. Raul de la Torre & A. M. Marmont. Multivariate analysis of risk factors for survival and relapse in chronic granulocytic leukaemia following allogeneic marrow transplantation: impact of disease related variables (Sokal score). *Bone Marrow Transplant* 1993: 12; 443-448

5.- J. Sierra, A. Graneda, J. Garcia, A. Valls, E. Carreras, M. Rovira, C. Canals, E. Martinez, C. Punti, M. Algara, P. Martin, A. Merino, M. J. Terol, A. Urbano-Ispizua & C. Rozman. Autologous bone marrow transplantation for acute leukemia: results and prognostic factors in 90 consecutive patients. *Bone Marrow Transplant* 1993: 12; 517-523

6.- J. L. Bidwell. Applications of the polymerase chain reaction to HLA class II typing. *Vox Sang* 1992: 63; 81-89

7.- P. Parham, E. Adams & k. L. Arnett. The Origins of HLA-A, B, C Polymorphism. *Immunological Reviews* 1995: 143; 141-180

8.- P. Krausa, M. Brywka III, D. Savage, K. M. Hui, M. Bunce, J. L. F. Ngai, D. L. T. Teo, Y. W. Ong, D. Barouch, C. E. M. Allsop, A. V. S. Hill, A. J. McMichael, J. G. Bodmer & M. J. Browning. Genetic polymorphism within HLA-A*02: significant allelic variation revealed in different populations. *Tissue Antigens* 1995: 45; 223-231

9.- P. Krausa, J. G. Bodmer, M. J. Browning. Defining the

common subtypes of HLA A9, A10, A28 and A19 by use of ARMS/PCR. *Tissue Antigens* 1993: 42; 91-99

10.- C. W. Summers, V. J. Hampson & G. M. Taylor. HLA class I non-coding nucleotide sequences, 1992. *European Journal of Immunogenetics* 1993: 20; 201-240

11.- S. Hoshino, A. Kimura, Y. Fukuda, K. Dohi & T. Sasazuki. Polymerase chain reaction-single-strand conformation polymorphism analysis of polymorphism in DPA1 and DPB1 genes: A simple, economical, and rapid method for histocompatibility testing. *Human Immunology* 1992: 33; 98-107

12.- M. Orita, Y. Suzuki, T. Sekiya & K. Hayashi. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 1989: 5; 874-879

13.- Y. Suzuki, T. Sekiya & K. Hayashi. Allele-specific polymerase chain reaction: A method for amplification and sequence determination of a single component among a mixture of sequence variants. *Analytical Biochemistry* 1991: 192; 82-84

14.- R. Blasczyk, U. Hahn, J. Wehling, D. Huhn & A. Salama. Complete subtyping of the HLA-A locus by sequence-specific amplification followed by direct sequencing or single-strand conformation polymorphism analysis. *Tissue Antigens* 1995: 46; 86-95

15.- R. Blasczyk, J. Wehling, B. S. Kubens, U. Hahn, D. Huhn & A. Salama. A novel HLA-A24 allele (A*2405) identified by single-strand conformation polymorphism analysis and confirmed by solid-phase sequencing and isoelectric focusing. *Tissue Antigens* 1995: 46; 54-58

16.- T. M. Clay, J. L. Bidwell, M. R. Howard & B. A. Bradley.

PCR-fingerprinting for selection of HLA matched unrelated marrow donors. *Lancet* 1991: 337; 1049-52

17.- A. Bhattacharyya & D. M. J. Lilley. The contrasting structures of mismatched DNA sequences containing looped-out bases (bulges) and multiple mismatches (bubbles). *Nucleic Acids Res* 1989: 17; 6821-6840

18.- F. Aboul-ela, D. Koh & I. Tinoco. Base-base mismatches. Thermodynamics of double helix formation for dCA3XA3G + dCT3YT3G (X, Y = A, C, G, T) *Nucleic Acids Res* 1985: 13; 4811-4824

19.- J. Y. Tong, A. Hammad, W. A. Rudert, M. Trucco & S. Hsia. Heteroduplexes for HLA DQB1 identity of family members and kidney donor-recipient pairs. *Transplantation* 1994: 57; 741-745

20.- R. Sorrentino, I. Cascino & R. Tosi. Subgrouping of DR4 alleles by DNA heteroduplex analysis. *Human Immunology* 1992: 33; 18-23

21.- M. White, M. Carvalho, D. Derse, S. O'Brien & M. Dean. Detecting single base substitutions as heteroduplex polymorphisms. *Genomics* 1992: 12; 301-306

22.- M. Carrington, M. White, M. Dean, D. Mann & F. E. Ward. The use of DNA heteroduplex patterns to map recombination within the HLA class II region. *Human Immunology* 1992: 33; 114-121

23.- N. A. P. Wood, T. M. Clay & J. L. Bidwell. HLA-DR/Dw matching by PCR Fingerprinting: The origin of PCR fingerprints and further applications. *European Journal of Immunogenetics* 1991: 18; 147-153

24.- M. Allen, L. Liu & U. Gyllensten. A comprehensive

polymerase chain reaction-oligonucleotide typing system for HLA class I A locus. *Human Immunology* 1994: 40; 25-32

25.- X. Gao, I. B. Jakobsen & S. W. Serjeantson. Characterization of the HLA-A polymorphism by locus-specific polymerase chain reaction amplification and oligonucleotide hydridization. *Human Immunology* 1994: 41; 267-279

26.- A. Selvakumar, C. B. Granja, M. Salazar, S. M. Alosco, E. J. Yunis & B. Dupont. A novel subtype of A2 (A*0217) isolated from the South American Indian B-cell line AMALA. *Tissue Antigens* 1995: 45; 343-347

27.- E. W. Petersdorf & J. A. Hansen. A comprehensive approach for typing the alleles of the HLA-B locus by automated sequencing. *Tissue Antigens* 1995: 46; 73-85

28.- N. Cereb, P. Maye, S. Lee, Y. Kong & S.Y. Yan. Locus specific amplification of HLA class I genes from genomic DNA. Tissue Antigens 1995.

29.- J Lyons. Analysis of *ras* gene point mutations by PCR and oligonucleotide hybridisation. In eds MA Innis et al, PCR Protocols, a guide to methods and applications.Academic Press, 1990.

30.- GE Hawes, L Struyk, PJ van den Elsen. Differential usage of T cell receptor V gene segments in $CD4^+$ and $CD8^+$ subsets of T lymphocytes in monozygotic twins. *J Immunol* 1993: 150; 2033-2045.

31.- C Giachino, MP Rocci, G De Libero, G Oderda, N Ansaldi and N Migone. An alternative approach to the assessment of T-cell clonality in celiac disease intestinal lesions through cDNA heteroduplex analysis of T-cell receptor VJ junctions. *Human Immunology* 1994: 40; 303-311.

CLAIMS

1. A method for separating a DNA molecule from a mixture of DNA molecules having the same number of nucleotides but different base sequences, which method comprises:

(i) amplifying the DNA molecules in the mixture;

(ii) hybridising single strands of the amplified DNA molecules with a complementary strand of a reference DNA molecule so as to form duplexes; and

(iii) separating the duplexes.

2. A method according to claim 1 which comprises

(i) amplifying the DNA molecules in the mixture employing a pair of primers in which one of the primers carries a ligand, so as to produce an amplified mixture of double-stranded DNA molecules in which one of the strands carries a ligand;

(ii) contacting the amplified mixture of double-stranded DNA molecules with a receptor on a solid support under conditions such that the ligand binds to the receptor;

(iii) separating the mixture of double-stranded DNA molecules into single-strands and removing the strands that are not bound to the support by the ligand;

(iv) recovering the remaining strands from the support;

(v) mixing the recovered strands with a complementary strand of a reference DNA molecule so as to form duplexes; and

(vi) separating the duplexes.

3. A method according to claim 2 wherein the complementary strand of the reference DNA molecule is provided by

(i) amplifying the reference DNA molecule employing a pair of primers in which one of the primers carries a ligand, so as to produce amplified double-stranded reference DNA molecule in which one of the strands carries a ligand;

(ii) contacting the double-stranded reference DNA molecule with a receptor on a solid support under conditions such that the ligand binds to the receptor;

(iii) separating the double-stranded reference DNA molecule into single-strands and removing the strand that is not bound to the support by the ligand; and

(iv) recovering the remaining strand from the support.

4. A method according to claim 1 which comprises

(i) amplifying the DNA molecules in the mixture employing a pair of primers in which one of the primers carries a high molecular weight molecule, so as to produce an amplified mixture of double-stranded DNA molecules in which one of the strands carries a high molecular weight molecule;

(ii) separating the mixture of double-stranded DNA molecules into single strands;

(iii) mixing the single strands with a complementary strand of a reference DNA molecule so as to form duplexes;

and

(iv) separating the duplexes.

5. A method according to claim 4 wherein the complementary strand of the reference DNA molecule is provided by

(i) amplifying the reference DNA molecule employing a pair of primers in which one of the primers carries a high molecular weight molecule, so as to produce an amplified double-stranded reference DNA molecule in which one of the strands carries a high molecular weight molecule; and

(ii) separating the double-stranded reference DNA molecule into single strands.

6. A method according to claim 1 which comprises

(i) amplifying a single strand of each of the DNA molecules in the mixture;

(ii) mixing the amplified single strands with a complementary strand of a reference DNA molecule so as to form duplexes; and

(iii) separating the duplexes.

7. A method according to claim 6 wherein the complementary strand of the reference DNA molecule is provided by amplifying a single strand of the reference DNA molecule.

8. A method according to any one of the preceding claims wherein the reference DNA molecule has a known sequence.

9. A method according to any one of the preceding claims wherein the duplexes are separated by gel electrophoresis.

10. A method according to claim 9 wherein the electrophoresis is performed under denaturing conditions.

11. A method for identifying a DNA molecule in a mixture of DNA molecules having the same number of nucleotides but different base sequences, which method comprises separating the DNA molecules by a method as defined in claim 9 or 10, and comparing the positions of the separated duplexes on the gel with the position of a control DNA molecule.

12. A method for identifying a DNA molecule in a mixture of DNA molecules having the same number of nucleotides but different base sequences, which method comprises separating the DNA molecules by a method as defined in any one of claims 1 to 10 and sequencing each of the separated DNA molecules, carrying out sequence specific primer (SSP) amplification analysis or carrying out sequence specific oligonucleotide (SSO) analysis.

13. A method according to any one of the preceding claims wherein the mixture of DNA molecules is a mixture of alleles of a polyallelic gene.

14. A method according to claim 13 wherein a reference allele is used which has the same serotype as at least one of the alleles in the mixture of alleles.

15. A method according to claim 13 or 14 wherein the mixture of alleles is from a prospective recipient or a prospective donor in a tissue or organ transplant operation.

16. A method for determining whether a prospective recipient in a tissue or organ transplant operation has

alleles of a gene that are compatible with the alleles of a prospective donor in the operation, which method comprises

(i) amplifying the alleles of the prospective recipient employing a pair of primers in which one of the primers carries a ligand, so as to produce amplified double-stranded alleles of the prospective recipient in which one of the strands carries a ligand;

(ii) contacting the amplified double-stranded alleles with a receptor on a solid support under conditions such that the ligand binds to the receptor;

(iii) separating the double-stranded alleles into single-strands and removing the strands that are not bound to the support by the ligand;

(iv) recovering the remaining strands from the support;

(v) mixing the recovered strands with complementary strands of the alleles of the prospective donor so as to form test duplexes;

(vi) separating the test duplexes by gel electrophoresis; and carrying out one or more of the following steps:

(vii) comparing the positions to which the test duplexes migrate on the gel with the position of a control DNA molecule;

(viii) sequencing the test duplexes;

(ix) sequence specific primer (SSP) amplification analysis; and

(x) sequence specific oligonucleotide (SSO) analysis.

17. A method according to claim 16 wherein the complementary strands of the alleles of the prospective donor are provided by

(i) amplifying the alleles of the prospective donor employing a pair of primers in which one of the primers carries a ligand, so as to produce amplified double-stranded alleles of the prospective donor in which one of the strands carries a ligand;

(ii) contacting the amplified double-stranded alleles with a receptor on a solid support under conditions such that the ligand binds to the receptor;

(iii) separating the double-stranded alleles into single-strands and removing the strands that are not bound to the support by the ligand; and

(iv) recovering the remaining strands from the support.

18. A method according to claim 16 or 17 wherein the prospective donor is selected to have alleles of the same serotype as the prospective recipient.

19. A method according to any one of claims 16 to 18 wherein the control DNA is a homoduplex between two strands of the same allele and migration of the test duplexes to the same position on the gel as the homoduplex indicates that the prospective recipient and the prospective donor have the same alleles.

20. A method according to any one of claims 13 to 19 wherein the alleles are of a human leucocyte antigen (HLA) class I gene or an HLA class II gene.

21. A method according to claim 20 wherein the alleles are of the HLA-A gene, the HLA-B gene or the HLA-C gene.

22. A method according to any one of claims 2, 3 and 9 to 21 wherein the ligand is biotin and the receptor is streptavidin.

23. A method according to any one of claims 2, 3 and 9 to 22 wherein the solid support is magnetic beads, and the strands which do not carry a ligand are removed by attracting the beads to a magnet and washing the beads under conditions such that the double-stranded DNA molecules dissociate into single strands.

24. A method according to any one of the preceding claims wherein amplification of the DNA molecules is performed by polymerase chain reaction (PCR).

25. A method according to any one of claims 15 to 24 wherein the prospective recipient and the prospective donor are a prospective bone marrow recipient and a prospective bone marrow donor, or a prospective kidney recipient and a prospective kidney donor.

ABSTRACT

METHODS FOR SEPARATING AND IDENTIFYING DNA MOLECULES

The invention provides a method for separating a DNA molecule from a mixture of DNA molecules having the same number of nucleotides but different base sequences, which method comprises:

(i) amplifying the DNA molecules in the mixture;

(ii) hybridising single strands of the amplified DNA molecules with a complementary strand of a reference DNA molecule so as to form duplexes; and

(iii) separating the duplexes.

The different DNA molecules in the original mixture give rise to duplexes having different numbers, positions or types of mismatches. This allows the duplexes to be separated by, for example, gel electrophoresis. The separated duplexes can then be analysed to identify the DNA molecules that were present in the original mixture.

FIGURE 1

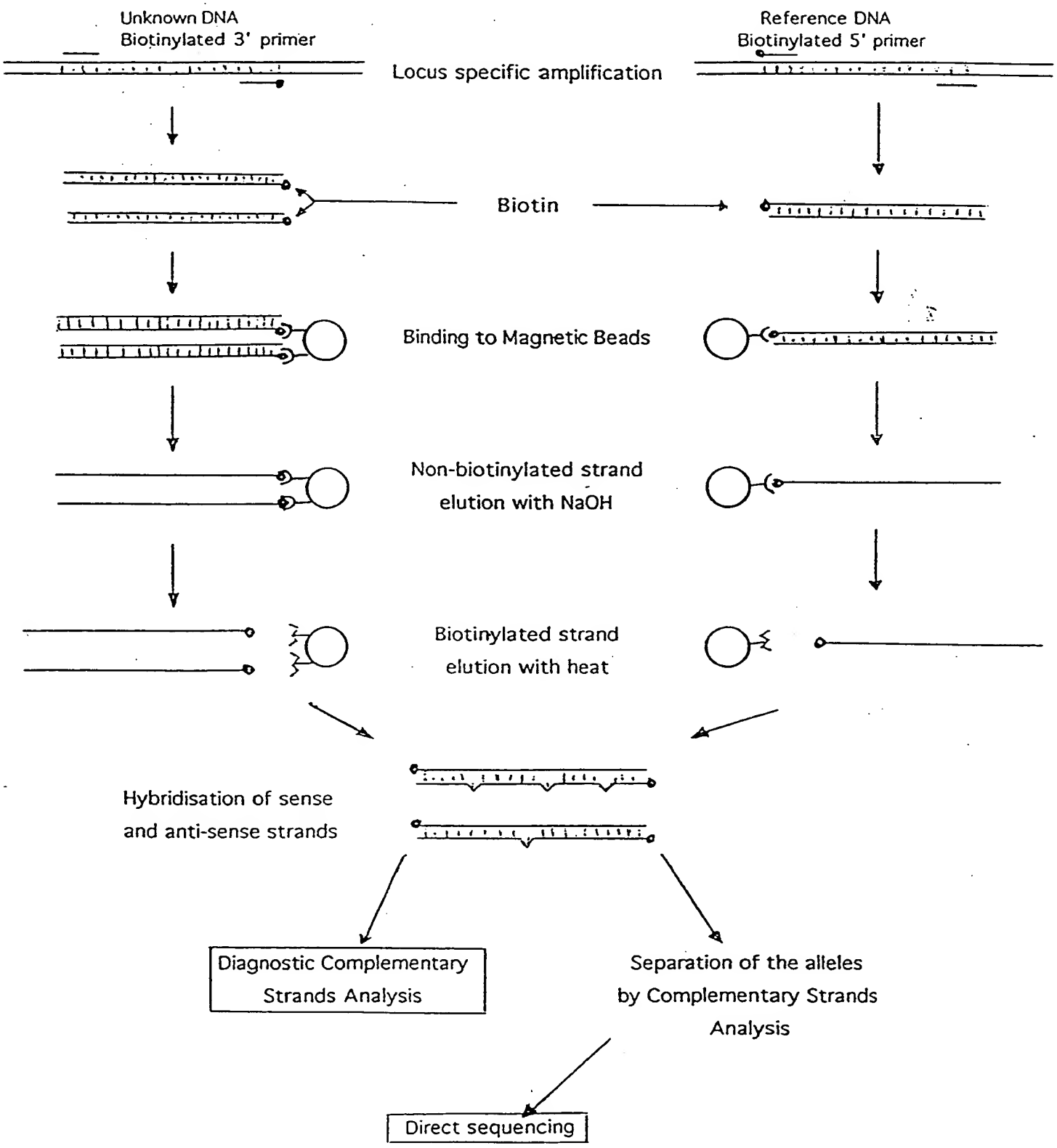

Unknown DNA
Biotinylated 3' primer
Reference DNA
Biotinylated 5' primer
Locus specific amplification
Biotin
Binding to Magnetic Beads
Non-biotinylated strand
elution with NaOH
Biotinylated strand
elution with heat
Hybridisation of sense
and anti-sense strands
Diagnostic Complementary
Strands Analysis
Separation of the alleles
by Complementary Strands
Analysis
Direct sequencing

FIGURE 2

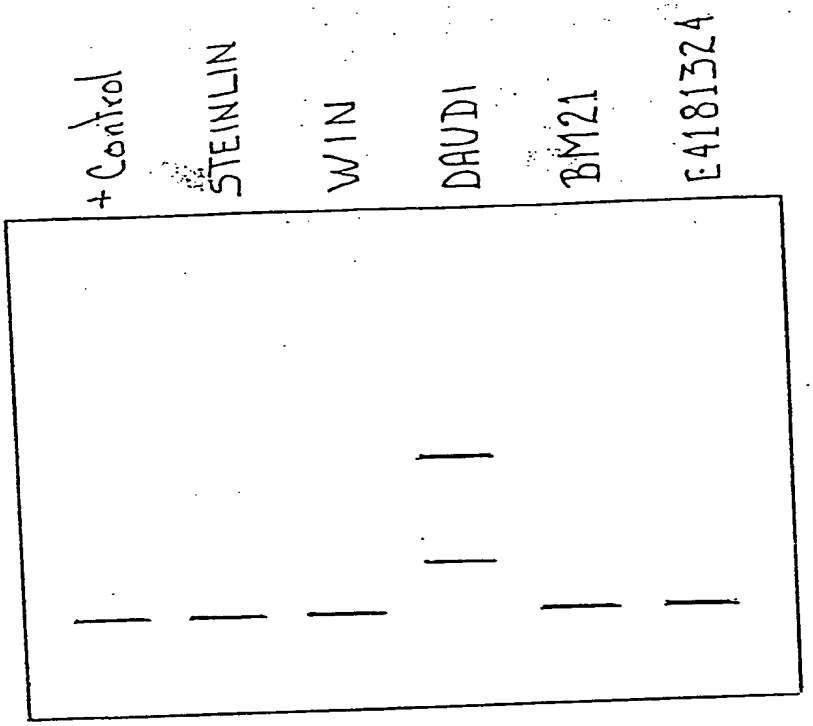


FIGURE 3

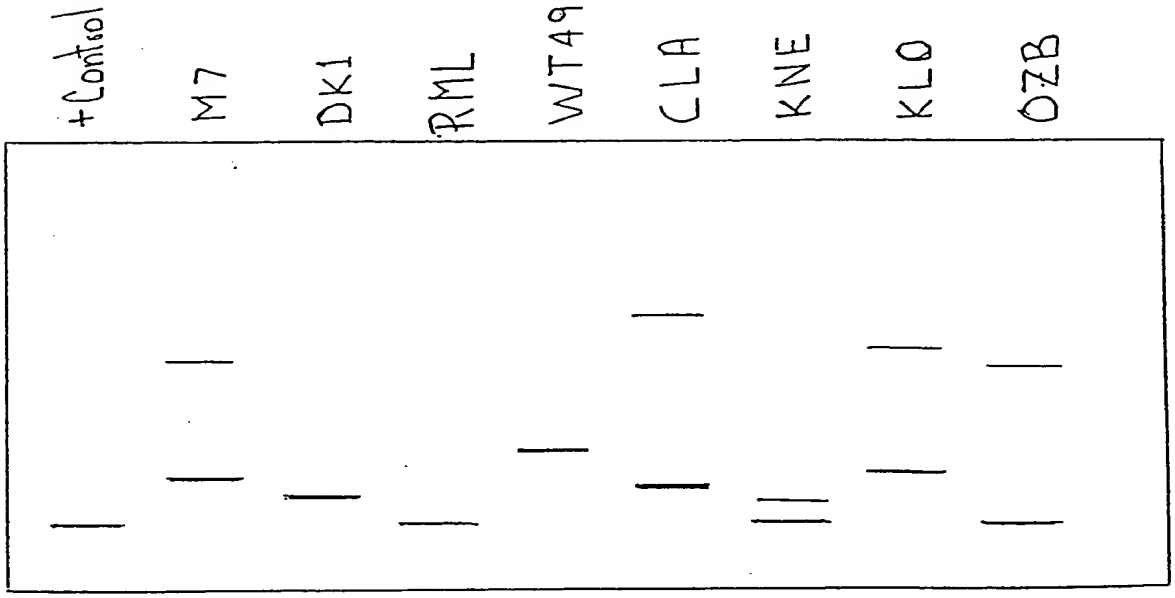


FIGURE 4

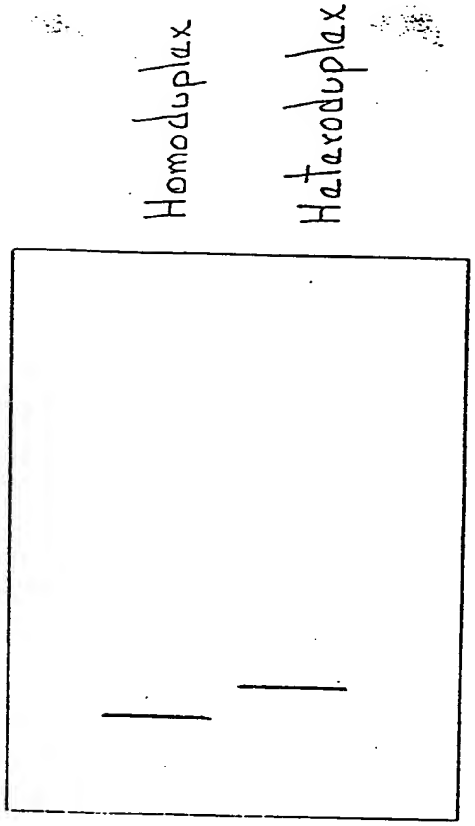
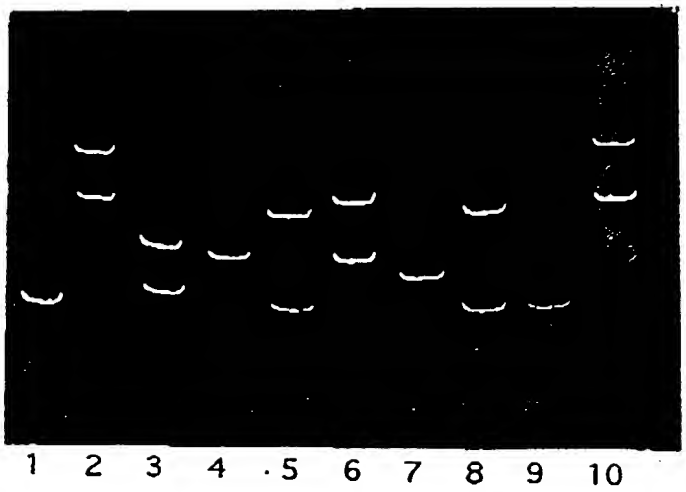


FIGURE 5



PCT/GB 96/02960 [redacted] 29.11.96 - JA KEMP + [redacted]

96003041